An effect of age on implicit memory that is not due to explicit contamination: Implications for single and multiple-systems theories

Emma V. Ward, Christopher, J. Berry, and David R. Shanks

University College London

Author Note

Emma V. Ward, Christopher, J. Berry, and David R. Shanks, Department of Cognitive, Perceptual and Brain Sciences, University College London, England.

Emma V. Ward is now at the Department of Psychology, University of York, and Christopher J. Berry is at the School of Psychology, Plymouth University.

This research was supported by an ESRC studentship award. We would like to thank Qizhang Sun for help with data collection in Experiments 3b and 3c.

Correspondence concerning this article should be addressed to Emma V. Ward, Department of Psychology, University of York, Heslington, York, YO10 5DD. E-mail: emma.ward@york.ac.uk

Abstract

Recognition memory is typically weaker in healthy older relative to young adults, while performance on implicit tests (e.g., repetition priming) is often comparable between groups. Such observations are commonly taken as evidence for independent explicit and implicit memory systems. On a picture version of the continuous identification with recognition (CID-R) task, we found a reliable age-related reduction in recognition memory, while the age effect on priming did not reach statistical significance (Experiment 1). This pattern was consistent with the predictions of a formal single-system model. Experiment 2 replicated these observations using separate priming (CID) and recognition phases, while a combined data analysis revealed a significant effect of age on priming. In Experiment 3 we provide evidence that priming in this task is unaffected by explicit processing, and we conclude that the age difference in priming is unlikely to have been driven by differences in explicit processing between groups of young and older adults (‘explicit contamination’). The results support the view that explicit and implicit expressions of memory are driven by a single underlying memory system.

Keywords: aging, priming, implicit memory, recognition, explicit contamination

Memory can be measured directly or indirectly. Direct or *explicit* tests (e.g., recognition) require deliberate recollection of specific information from a prior study episode, whereas indirect or *implicit* tests measure memory of previously studied information in a seemingly unrelated task (e.g., perceptual identification). Priming is a commonly used index of implicit memory. It refers to a long-term change in behavioural response to an item as a result of prior exposure to it, and usually takes the form of facilitated processing. For example, previously studied words or pictures are usually identified more quickly than new ones.

The question of whether there are distinct memory systems driving explicit and implicit memory has provoked extensive research over the past few decades. Performance on explicit and implicit tests has been shown to dissociate under numerous experimental manipulations (reviewed in Roediger & McDermott, 1993) and a common interpretation is that this reflects processing in independent cognitive systems (e.g., Gabrieli, 1998; Squire, 1994, 2004, 2009; Tulving & Schacter, 1990). The dissociation in normal aging has provided particularly compelling evidence for this ‘multiple-systems’ account – older individuals typically show decrements on explicit tests relative to healthy young individuals despite comparable priming levels (reviewed in Fleischman, 2007). The typical interpretation of this finding is that an explicit (or declarative) system is affected by age whereas an implicit (or nondeclarative) system is not.

Age-invariant priming has been reported on tests of word-stem completion (e.g., Light & Singh, 1987; Park & Shaw, 1992), word identification (e.g., Light, La Voie, Valencia-Laver, Albertson-Owens, & Mead, 1992; Light & Singh, 1987), picture naming (e.g., Mitchell, Brown, & Murphy, 1990; Sullivan, Faust, & Balota, 1995), degraded picture naming (e.g., Russo & Parkin, 1993), and object decision (e.g., Schacter, Cooper & Valdiserri, 1992; Soldan, Hilton, Cooper, & Stern, 2009). However, other studies have reported priming deficits for older individuals on some of the same tests (e.g., Abbenhuis, Raaijmakers, Raaijmakers, & Van Woerden, 1990; Chiarello & Hoyer, 1988; Hultsch, Mason, & Small, 1991). These results favor the alternative view, the single-system account, which states that explicit and implicit memory are driven by a common underlying system. It has been suggested that the discrepancies between studies are largely due to methodological differences (e.g., Fleischman, 2007), but it remains unclear whether priming is truly affected by aging.

How convincing are prior reports of spared priming in old age? Most published demonstrations of age-invariant priming rest on a null result (a failure to find a statistically significant difference between groups), and it is possible that many studies lacked the required statistical power to detect a significant priming difference. Indeed, priming is usually numerically lower in older relative to young adults, and a meta-analysis by La Voie and Light (1994) showed a small but significant effect of age on priming. Furthermore, indirect memory tests are usually associated with lower reliability levels than direct tests, and are thus less likely to be sensitive to age differences (e.g., Buchner & Wippich, 2000). One must therefore question the necessity of drawing a distinction between explicit and implicit memory systems based on evidence of a statistically significant group difference on a measure of explicit memory coupled with a non-significant priming difference.

Prior conclusions have also been limited because samples of explicit and implicit memory have traditionally been measured in separate experimental phases. Scores may dissociate because there is a longer study-test delay for one task than the other, and/or because participants adopt different response strategies or levels of motivation in the two tasks when they are presented separately. For samples of explicit and implicit memory to be truly comparable, they need to be taken for the same items at around the same point in time (see Stark & McClelland, 2000). Dissociations produced under these circumstances constitute more persuasive evidence for multiple memory systems relative to when items are judged on two separate occasions.

Computational models can offer considerable theoretical insights regarding empirical dissociations. Formal single-system models have successfully reproduced several dissociations that have previously been taken as support for multiple memory systems (e.g., Berry, Shanks, & Henson, 2006; Berry, Shanks, & Henson, 2008a; Berry, Shanks, & Henson 2008b; Berry, Shanks, Li, Rains, & Henson, 2010; Berry, Shanks, Speekenbrink, & Henson, 2012; Kinder & Shanks 2001; 2003; Shanks & Perruchet, 2002; Shanks, Wilkinson, & Channon, 2003). The model by Berry and colleagues assumes that a single memory signal drives performance on explicit and implicit tasks, but that there are independent sources of random noise, the variance of which is greater in the implicit task (an assumption which is fortified by the generally lower reliability levels found in implicit relative to explicit tests; Buchner & Wippich, 2000). The model has reproduced dissociations between recognition and priming, such as those generated by manipulating attention at study (e.g., Butler & Klein, 2009), and those seen in individuals with amnesia due to damage to the hippocampus or medial temporal lobe (e.g., Conroy, Hopkins, & Squire, 2005; see Berry et al., 2008b, 2010, 2012).

Until recently there have been few attempts to test formal multiple-systems models. Berry et al. (2012) developed two such models in which two independent signals either make unique contributions to explicit and implicit memory or are assumed to have some degree of correlation. Not only did the single-system model reproduce the qualitative dissociation observed in amnesia in the Conroy et al. (2005) study, but model selection on the basis of the Akaike Information Criterion indicated that it fit the data better than the multiple-systems models. Thus, many empirical observations which on the surface appear to be indicative of multiple systems are not incompatible with the single-system view.

More robust evidence for multiple-systems would come from a demonstration that priming is completely preserved in old age (i.e., equivalent in young and older adults), despite a clear group difference in recognition memory when the two are measured consecutively on each test trial. The present study aimed to establish whether such a pattern can be produced. This study is the first to compare recognition and priming in young and older individuals using the continuous identification with recognition (CID-R) paradigm (e.g., Conroy et al., 2005; Stark & McClelland, 2000) in which the two are measured concurrently on each test trial. Each trial comprised a speeded masked picture identification (priming measure) followed immediately by a yes (studied)/no (nonstudied) recognition judgement for the item. Recognition memory is typically affected by age, thus the study allowed us to test the competing predictions of the multiple and single-system perspectives regarding the effect on priming. Because the multiple-systems account assumes that an implicit memory system (which operates independently of an explicit system) is resistant to the effects of aging, it predicts no effect of age on priming. On the other hand, the single-system view predicts an age effect on priming that is smaller than that on recognition.

Experiment 1

We compared the performance of young and older adults on the CID-R task. Participants studied pictures of everyday objects both immediately and 60 minutes prior to the test. The delay was included to reduce the strength of memory for a subset of items as much as possible. We anticipated reductions in recognition memory as a function of age and delay, and the question of whether priming would be similarly affected was of primary interest.

Method

Participants

Twenty young (seven male) and 20 older (two male) adults participated for a small payment. The young adults were students from University College London (UCL), recruited through an advertisement on an internal website. The older adults were members of the University of the Third Age (U3A) organisation. All were native English speakers who reported good health. Participant demographic information is summarised in Table 1.[1]

Materials

The stimuli were pictures of everyday objects from ten categories: animals, clothing, fruit and vegetables, electrical appliances, musical instruments, transportation vehicles, kitchen utensils, insects, tools, and furniture. Items within these categories were selected based on the category norms collected by Van Overschelde, Rawson, and Dunlosky (2004). There were 120 critical items, each depicting a colour photograph of an object on a 400 x 400 pixel white background (Figure 1A). Sixty pictures, six from each object-category, appeared in the study phases (30 in each phase) and were also presented at test to serve as old (studied) items. The other 60 pictures served as new items and were only presented at test. Four sets of 30 pictures were rotated between participants to serve as old/new items, and to appear in the initial/delayed study phase. The mask used in the identification task was a 400 x 400 pixel grid randomly filled with fragments of non-critical item pictures.

Design and procedure

The experiment used a mixed factorial design with the between-subjects factor age group (young/older adult) and the within-subjects factor delay (60 min/no delay).

The experimental procedure, identical for both groups, consisted of four parts: an initial study phase, a 60 minute interval, a second study phase, and finally the CID-R test. Participants were tested individually, and the duration of the experiment was approximately 90 minutes. In this and subsequent experiments, the task was programmed in Matlab 6 using the Cogent 2000 toolbox, and administered on a PC with a screen resolution of 1024 x 768 pixels. Viewing distance was approximately 50 cm.

Initial study. Participants performed an incidental encoding task which involved matching briefly presented pictures to object categories. Each trial was presented as follows: (1) a fixation point (‘+’) was presented in centre screen for 500 ms, (2) a picture (e.g., a dog) was presented for 250 ms, (3) The instruction “Which category was the object from?” appeared at the top centre of the screen, and two category options were displayed (e.g., F =

animal / J = musical instrument). Participants were instructed to use the ‘F’ and ‘J’ keyboard keys to select the correct option. No time limit was imposed on participants to respond, and the choice categories remained on the screen until a keypress was made. (4) Finally, there was a 1000 ms blank screen prior to the start of the next trial. There were 30 randomised trials in total, plus 5 practice trials.

Interval testing. The interval between the initial and second study phases was 60 minutes. The following battery of tests was administered: (1) Demographic and health questionnaire, (2) Near Vision Test, (3) Wechsler Test of Adult Reading, (4) WAIS-III subtests: Vocabulary and Digit Symbol Substitution, (5) Mini Mental State Exam (older adults only). Breaks were provided.

Second study. Next, participants performed the second study phase which was identical in format to the initial study task, but comprised a different set of 30 critical items.

CID-R test. Immediately following the second study phase, participants were given instructions for the CID-R task. Participants were not informed of this task in advance. Each trial consisted of a speeded masked-picture identification in which response times (RT) were measured, and a recognition judgement. Each trial was self-initiated by the participant, and began with the identification task: A mask was initially presented in centre screen for 500 ms. A picture (old or new) was then presented for 17 ms, followed immediately by the mask for 233 ms (making a 250 ms block). These block presentations were repeated, with the duration of picture presentations increasing by 17 ms on every alternate block while the total block duration remained constant, thus making the picture gradually more visible (Figure 1B). Participants were required to identify the picture as quickly and accurately as possible, pressing the ‘Enter’ key when they knew the identity of the object. RTs were captured on the keypress, at which point the picture disappeared and participants were prompted to type the object name into a box. The block presentations ceased at 7 sec (30 blocks) after initiation if identification had not taken place, and any such trials were discarded.

The recognition segment of the trial immediately followed each identification. The picture was presented once more and participants were required to make a judgement as to whether they thought it was shown in either of the study phases using a 6-point scale where 1 = *very sure no*; 2 = *fairly sure no*; 3 = *guess no*; 4 = *guess yes*; 5 = *fairly sure yes*; 6 = *very sure yes*. No time limit was imposed in making recognition judgements, and no feedback was provided. There were 120 randomised trials in total – 60 old (30 items from each study phase) and 60 new.

Results

In this and subsequent experiments, an alpha level of .05 was used, and all *t*-tests were two-tailed unless otherwise stated.

Study phases

Mean categorization accuracy was at 98.3% (*SD* = 2.29) for young and 96.5% (*SD* = 4.77) for older adults on the initial study phase, and 97.7% (*SD* = 3.26) for young adults and 98.5% (*SD* = 2.53) for older adults on the second study phase. Performance was not statistically different between groups in either phase: $t(38) = 1.55$, $p = .13$, and $t = 0.90$, $p = .37$, for the initial and second study phases, respectively. In this and subsequent experiments, items associated with incorrect study phase responses were removed from further analysis.

Recognition

Ratings 4-6 (‘*yes*’ – old) and 1-3 (‘*no*’ – new) on the 6-point scale were collapsed. For each participant, the proportion of hits (old pictures judged old) and false alarms (new pictures judged old) were used to calculate d' (Figure 2A).[2]

Discrimination was significantly greater than chance (i.e., $d' > 0$) for young and older adults, for items studied immediately prior to testing and those studied 60 min prior to testing (all *t*’s(19) > 7, *p*’s < .001, *d*’s > 1.72). There was a significant main effect of age group, $F(1, 38) = 4.26$, $p = .04$, $\eta_p^2 = .10$, and delay, $F(1, 38)$, = 5.76, $p = .02$, $\eta_p^2 = .13$, and no significant interaction between the factors, $F(1, 38) = 0.16$, $p = .69$. Young adults' recognition performance was superior to that of older adults for items studied immediately prior to testing, $t(38) = 2.03$, $p = .05$, $d = 0.64$, and items studied 60 min prior to testing, $t(38) = 1.78$, $p = .04$, $d = 0.56$ (one-tailed). Recognition was significantly reduced for items studied 60 min prior to testing relative to those studied immediately before for young adults, $t(19) = 1.67$, $p = .05$, $d = 0.32$, and older adults, $t(19) = 1.83$, $p = .04$, $d = 0.32$ (both one-tailed).

Priming

The following steps were performed on each participant's raw RT data (Table 2) to obtain a priming score: (1) trials associated with incorrect identifications were removed; (2) RTs for old and new items were averaged separately and trials with identification latencies faster than 200 ms or greater than 3 SD from the mean were removed; (3) Priming was then calculated as the difference in the median RT between new and old items, expressed in proportion to the individual's baseline (new item) RT.[3] Priming scores were averaged within each group (Figure 3A).

Priming was strong and significantly above chance (i.e., > 0 ms) for items studied immediately prior to testing ($t(19) = 3.68$, $p = .002$, and $t(19) = 2.79$, $p = .01$, for young and older adults, respectively), and items studied 60 min prior to testing ($t(19) = 4.60$, $p < .001$, and $t(19) = 2.78$, $p = .01$, for young and older adults, respectively, all d's > 0.62). However, a repeated measures ANOVA with the between-subjects factor age group revealed no significant main effect of age group, $F(1, 38) = 1.78$, $p = .19$, or delay, $F(1, 38) = 0.02$, $p = .89$, and no significant interaction, $F(1, 38) = 0.04$, $p = .85$.

Discussion

Recognition memory was reduced by age and delayed testing, while priming was not significantly affected. There was, however, a clear numerical trend – priming was lower in older adults compared to young, so we are reluctant to conclude that priming is preserved in old age. There are two possible explanations for this pattern (other than it being due to sampling error): First, there is a genuine decline in priming with age which this experiment failed to detect statistically, and secondly, priming was reduced in older individuals compared to young due to explicit contamination in the priming (CID) task.

Participants may use explicit processing in an implicit test if they become aware that some items were previously studied (termed *test awareness* or *awareness*), and such a strategy is said to ‘contaminate’ the priming measure. In the CID-R paradigm, participants are made aware at test that some items were shown at study, so they may attempt to retrieve study items from memory in an effort to facilitate their identifications of the objects. If such a strategy can affect priming (e.g., by reducing identification speeds of old items, and/or increasing identification speeds of new items), it is likely to be of a greater benefit to the performance of young individuals, due to their superior explicit memory. That is, explicit processing is more likely to result in boosted priming in young individuals (see Geraci & Barnhardt, 2010; Mitchell & Bruss, 2003; Russo & Parkin, 1993). To examine whether explicit processing could have mediated the age differential pattern in priming, in Experiment 2 we once again compared the performance of young and older adults, but this time we attempted to reduce the likelihood that participants would engage in explicit processing in the priming task. To preview, after observing the same pattern of results when the likelihood of explicit processing was reduced, we turn to the prediction that low statistical power in these experiments masked a small but genuine decline in priming with age.

Experiment 2

All aspects of the design and procedure were the same as in Experiment 1, except that the CID and recognition tasks were presented separately, and we introduced a post-test awareness questionnaire. Thus, there were 6 distinct phases in this experiment: the initial study phase, a 60 minute interval, the second study phase, the CID task, the recognition task, and finally the awareness questionnaire. The purpose of separating the CID and recognition phases was to minimise test awareness (and thus the likelihood of explicit processing) in the priming phase. We assumed that participants would be less likely to become aware in the CID task when no reference is made to the prior study episode and when they are not required to make a recognition judgement after every identification trial. Observing a similar age-differential pattern of priming as in Experiment 1 would suggest that the effect is not due to differences between groups in the successful use of an explicit strategy.

Method

Participants

Eighteen young (seven male) and 18 older (two male) adults participated for payment of £5. The young adults were again recruited through the UCL participant database, and the older adults through the U3A. All were native English speakers who reported good health. Participant demographic information is summarised in Table 3.

Materials and procedure

The picture stimuli, categories and priming mask were the same as those used in Experiment 1, and the experimental procedure was the same, except that the CID and recognition tasks were presented separately as described above. All participants performed the CID task prior to the recognition task so as to maintain low levels of test awareness in the priming phase. Different critical items were presented in the CID and recognition tasks as we found ceiling recognition performance in a pilot study that used the same items in both phases. Thirty pictures from each study phase later served as old items in the CID task, and an additional 30 pictures which were included in the initial study phase served as old items in the recognition task, but were not presented in the CID phase. There were 60 new items in the CID task, and 30 in the recognition task.

At the end of the experiment participants completed a questionnaire to gauge their awareness during the CID phase. The questionnaire was similar to that introduced by Bowers and Schacter (1990), and included the following items: (1) *What do you think was the purpose of the identification task you performed?* (2) *Do you think that any of the pictures* *you identified were previously presented in the first parts of the experiment?* If participants failed to notice that any of the pictures were previously studied, they were classified as unaware and were not required to complete the rest of the questionnaire. If they noticed that some pictures were previously studied they were asked to complete the following questions: (3) *Were you aware that some of the pictures had been shown before as you were performing the task, or did you become aware of this afterwards/in hindsight?* (4) *Did you suspect prior to the start of the identification task that you would be tested on your memory of the pictures?* (5) *Did you try to use your memory of the pictures to help you in this task?* (6) *If yes, do you think this strategy helped you, and how so?* Participants who stated they became aware in hindsight were classified as unaware at the time of testing, and all other participants were deemed aware.

Results and Discussion

Study phases

Categorization accuracy was at 98.4% correct for both young and older adults on the initial study phase (young *SD* = 1.93; older *SD* = 2.25), and at 98.7% (*SD* = 2.60) for young, and 99.1% (*SD* = 1.92) for older adults on the second study phase. Performance was not statistically different between groups in either study phase (immediate study: $t(34) = 0.08$, $p = .94$; delayed study: $t(34) = 0.49$, $p = .63$.

Recognition

Recognition (Figure 2B) was significantly greater than chance for young, $t(17) = 7.52$, $p < .001$, and older adults, $t(17) = 7.32$, $p < .001$ (both *d*'s > 0.97). Discrimination (d') was superior in young relative to older adults, $t(34) = 3.41$, $p = .002$, $d = 1.19$.

Priming

Priming (Figure 3A; See Table 2 for RTs) was significantly greater than chance for items studied immediately prior to testing for young adults, $t(17) = 2.63$, $p = .02$, and older adults, $t(17) = 2.48$, $p = .02$, and items studied 60 min prior to testing for young adults, *t*(17)

= 2.20, p = .04 (all d's > 0.58). Priming for items studied 60 min before testing fell short of significance in the older adult group, $t(17) = 1.58$, p = .06 (one-tailed). An ANOVA revealed no significant main effect of age group, $F(1, 34) = 1.02$, p = .32, or delay, $F(1, 34) = 0.62$, p = .44, and no interaction, $F(1, 34) = 0.01$, p = .92.

Post-test awareness

Ten out of 18 young participants (55.5%), and 9/18 older participants (50%), were classified as aware during the CID phase. Collapsed across immediate and delayed items, priming did not significantly differ between aware (.12) and unaware (.13) young participants, $t(34) = 0.21$, p = .84, or aware (.08) and unaware (.08) older participants, $t(34)$ = 0.14, p = .89. There was also no significant difference between aware young versus older participants, $t(36) = 0.51$, p = .61, or unaware young versus older participants, $t(32) = 0.92$, p = .36. It should be noted that, because participants completed the questionnaire after the task was completed in its entirety, their recollection of the awareness experienced during the CID phase may have been distorted, and more participants may have actually been unaware during this phase than is reflected here.

In sum, we replicated the results of Experiment 1 – recognition was reliably reduced by age, and priming was numerically lowered. To test the possibility that the age difference in priming is genuine but failed to reach significance because of low power, we increased power by pooling and re-analysing the data from our experiments.

Re-analysis of pooled priming data

This analysis included the data from Experiments 1 and 2 as well as that of an additional experiment. The latter was almost identical to Experiment 1, hence only a brief outline is presented here. We included these data to increase statistical power as much as possible.[4] This experiment compared recognition and priming in 20 young and 20 healthy older adults (see Table 4 for details), however participants performed two CID-R tasks separated by 60 min following an initial study phase rather than a single CID-R task as in

Experiment 1. The priming and recognition data for this experiment are given in Table 5, along with a summary of the statistical outcomes. The pattern of results was identical to that reported previously: recognition but not priming was affected by age and delay.

Across experiments, we pooled the priming scores for young and older individuals for immediate and delayed items (n = 58 per group). Figure 3B illustrates the age difference in priming for immediate items. Priming was significantly above chance for young adults ($t(57) = 7.29$, $p < .001$, $d = 0.95$, $t(57) = 6.44$, $p < .001$, $d = 0.84$, for immediate and delayed items, respectively), and older adults ($t(57) = 5.64$, $p < .001$, $d = 0.74$, $t(57) = 5.87$, $p < .001$, $d = 0.76$, for immediate and delayed items, respectively). A 3 (experiment) x 2 (age group) x 2 (delay) ANOVA revealed significant main effects of experiment, $F(2, 110) = 7.17$, $p = .001$, $\eta_p^2 = .12$, and age group, $F(1, 110) = 4.10$, $p = .045$, $\eta_p^2 = .04$, but no effect of delay, $F(1, 110) = 0.46$, $p = .45$, and no interactions (largest $F = 1.59$, $p = .21$).

Critically, the analysis confirms a reliable reduction in priming as a function of age (46% reduction collapsed across immediate and delayed items). Although the effect size is small, the findings are of considerable importance, as noted in the Introduction. If levels of priming in young and older individuals were equivalent, this would strongly suggest that priming and recognition are driven by distinct memory systems. In contrast, the results are consistent with the single-system view which predicts an age effect on priming that is weaker than that on recognition due to the lower reliability of priming measures. We estimated the reliability of the recognition and priming tasks across experiments using split-half correlations (see Buchner & Wippich, 2000).

Measure reliability. We computed two scores for recognition (d') and priming (proportion) for each participant by extracting odd and even test trials. We expected the correlation between scores to be larger in recognition than priming, and this was observed in all cases (see Table 6). We also computed aggregate correlation coefficients for the recognition and priming scores, and in all cases the correlation was significantly larger for recognition than priming (z = 2.64, p = .008, and z = 2.54, p = .01, for Experiments 1 and 2, respectively, and z = 4.41, p < 001, for the additional experiment used in the pooled data-analysis).

It should be noted that these split-half correlations are merely reliability estimates, and are likely to be noisy due to the small samples. They serve to demonstrate that the priming index is statistically less reliable than the recognition measure, as predicted by the single-system model. It is noteworthy, however, that many of the correlations for the priming measure were very low or even negative, and one might wonder how this is possible in a measure that plainly – at the aggregate level – captures a meaningful and robust construct. If, as assumed by the model, priming consists of a fixed, memorial, component as well as non-memorial noise, then the pattern of data is not so surprising. For instance, the fixed component could be fairly similar across both trials and participants, yielding a strong aggregate priming effect. If at the same time the noise component is uncorrelated across trials and participants, then reliability measured across odd and even test trials will be low or even zero. It will be an important question for future research to analyse reliability data in other types of priming or implicit memory measures (see Buchner & Brandt, 2003, for one such analysis of data similar to those obtained here).

Model analysis. We fit the data of Experiment 1 to the Berry et al. single-system model in order to gain further insight into the age effect in priming.[5] Details of model fitting can be found in the appendix. The mean expected values for recognition and priming are shown in Figures 2A and 3A. The model captured the trends for recognition to be greater in young relative to the older adults, and for immediate relative to delayed items. Crucially, the predicted effect of age on priming was small: the model predicts an age difference of .041 for immediate items, and .026 for delayed items. The fits provide further evidence that the data are consistent with the single-system view. Furthermore, the model predicts a weak correlation between the recognition and priming scores (for a discussion see Berry et al., 2006), and the greatest observed correlation in the pooled data was for items studied immediately before testing in young adults, $r = .33$, $p = .01$.

These observations provide evidence that low measure reliability coupled with low statistical power can mask a genuine reduction in priming with age. This has substantial implications for prior studies which have reported evidence of spared priming in old age. However, the conclusion that priming is reduced by age rests crucially on eliminating the possibility that our priming measure was contaminated by explicit processing. As outlined previously, if explicit processing in the CID task affects priming, it is likely to be of a greater benefit to the performance of young adults, and thus could explain the age effect. Because test awareness is a necessary condition for the use of explicit processing, in Experiment 2 we attempted to reduce awareness in the priming task by separating the CID and recognition phases. The results were identical to Experiment 1 when participants were test aware and the CID and recognition tasks were presented concurrently trial-by-trial. It therefore seems unlikely that the reliable age difference in priming we report was driven by explicit contamination, but we nevertheless conducted further experiments to clarify whether explicit processing in the CID task can enhance priming in young individuals.

Experiments 3a and 3b

We attempted to create optimal conditions for the successful use of an explicit strategy in the CID task. Groups of Informed participants were told prior to each CID trial whether the next item to be identified was previously studied (cued ‘old’) or new (cued ‘new’), and encouraged them to use this information to help identify the objects. The cues enabled participants to search memory of the previously studied items on appropriate trials (i.e., those cued ‘old’ and not those cued ‘new’), and we were interested in the effect on identification speeds and priming in comparison to Uninformed participants who received no explicit information and were rigorously monitored for spontaneous test awareness. If explicit processing boosts performance in the CID task (e.g., by reducing identification times for old items), then we expect to see greater priming in Informed participants.

In Experiment 3b Informed participants were also given category information ahead of each old-item trial (e.g., ‘old – animal’). Thus these participants were guided to explicitly search memory of a particular (small) set of previously studied items, which arguably provides the best possible opportunity for them to produce a rapid identification time.

In Experiment 3a, we also varied the ratio of old to new test trials – half the participants were exposed to a high proportion of old trials (80% old and 20% new), and half to a low proportion (20% old and 80% new) with the assumption that this would create further differences in test awareness in the Uninformed group (for a similar manipulation see Jacoby, 1983). Thus, there were 4 conditions in total: Informed High, Informed Low, Uninformed High, and Uninformed Low, and we assumed that the Uninformed Low condition would be a suitable ‘unaware’ group against which to compare priming in the other groups. Importantly, an effect of explicit processing on priming would *not* manifest itself in a main effect of proportion. All participants in the Informed group – whether exposed to a high or low proportion of old to new trials – are test aware and equally capable of successfully using an explicit strategy, so an interaction between the factors whereby the effect of proportion is only significant in the Uninformed group (with lower priming in the low relative to the high condition) would be required to conclude that explicit processing affects priming.

Method

Participants and design

107 first year undergraduate students from UCL took part in Experiment 3a as part of a course requirement. There were 24 males and 83 females with an overall mean age of 18.7 (SD = 0.8). The experiment used a 2 (Informed vs. Uninformed) x 2 (High vs. Low proportion old trials) between-subjects design, and participants were randomly distributed among the conditions – Informed High (n = 26), Informed Low (n = 27), Uninformed High (n

= 27), Uninformed Low (n = 27). In Experiment 3b there were 32 participants (11 male) with an overall mean age of 22.7 (SD = 3.3). All were UCL students who participated for a small payment. Participants were equally divided into the Informed and Uninformed groups.

Materials and procedure

In Experiment 3a 120 critical pictures were presented at study and 150 in the CID phase. In the High conditions, all the pictures presented at study were shown again at test (old items), along with an additional 30 new items. In the Low conditions, 30 old items were presented at test, along with 120 new items. Five sets of 30 items were rotated among old/new status. In Experiment 3b there were 60 critical study trials and 120 CID trials (60 old and 60 new). There were six object categories in total (animals, clothing, electrical appliances, fruit and vegetables, kitchen utensils and furniture), thus in both conditions, ten items within each category were previously studied and ten were new. Two sets of items were rotated.

The procedure for the study and CID phases was the same as described previously, with the exception that, for participants in Informed groups, the word ‘old’ or ‘new’ was presented in centre screen for 2000 ms prior to the start of each trial and participants were instructed to try to use the cues to help them identify the objects. Informed participants in Experiment 3b were also given a category cue prompt on old trials (e.g., ‘old – animal’; on new trials participants were just prompted with the word ‘new’). Participants in Uninformed groups witnessed a fixation cross for 2000 ms prior to the start of each trial, and were not informed that some items were previously studied. They were given the awareness questionnaire described previously at the end of the experiment.

Results and Discussion

Study phase

Categorization accuracy ranged from 95.7% to 98.3%, and performance did not significantly differ between groups in either experiment (greatest t = 1.51, p = .14).

Priming and post-test awareness

Priming was significantly greater than chance in all groups (all *t*'s > 2, *p*'s < .05, all *d*'s > 0.46). A two-way ANOVA indicated no main effect of informing participants in Experiment 3a, $F(1, 103) = 0.11$, $p = .74$ (Figure 4A), and priming in Experiment 3b did not differ between groups, $t(30) = 0.26$, $p = .79$ (Figure 4B; see Table 7 for RTs). This provides compelling evidence that explicit processing does not affect priming, as informing participants which items were previously studied provides excellent conditions for the successful use of an explicit strategy. The main effect of varying the proportion of studied trials in Experiment 3a was significant, $F(1, 103) = 4.02$, $p = .05$, $\eta^2_p = .04$, but there was no interaction between the factors, $F(1, 103) = 0.01$, $p = .92$. It is likely that the effect of proportion was driven by something other than awareness. Both the Informed High and Informed Low groups were test aware, so an interaction between the factors whereby the High and Low conditions differed only within the Uninformed group is required to conclude that the effect is mediated by awareness. A comparison of priming in the Uninformed High and Uninformed Low conditions revealed no significant difference, $t(52) = 1.25$, $p = .22$.

Of the participants in the Uninformed groups in Experiment 3a, 25/27 (92.6%) in the High condition, and 14/27 (51.9%) in the Low condition were deemed test aware. Priming in these participants did not significantly differ to that in unaware participants (High: aware M = .09, unaware M = .11, $t(25) = 0.18$, $p = .86$; Low: aware M = .06, unaware M = .05, $t(25) = 0.33$, $p = .74$). Eight out of 16 (50%) participants in the Uninformed group in Experiment 3b were deemed aware, but, again, priming did not significantly differ between aware (.10) and unaware (.11) participants, $t(14) = 0.48$, $p = .64$.

Experiment 3c

Providing old/new status and category information about items yields no speeding of RTs or enhancement of priming relative to uninformed conditions, but a stronger test is to compare a correctly informed condition with a misinformed condition. If explicit processing can affect priming then we would expect correct cues (e.g., ‘old’ before an old picture) to induce particularly faster identifications relative to incorrect cues (e.g., ‘new’ before an old picture), in which participants would be positively discouraged from engaging in a search of explicit memory. As such, we compared identification times in the CID task in two groups of participants who received incorrect cues on a subset of trials (the majority of items were correctly cued – 90 out of 120 total – so as to ensure the overall validity of the cues). Participants were given an ‘old/new’ prompt before each trial as described previously, but one group (Misinformed New) received ‘old’ cues prior to 30 (out of 60 total) new-item trials, and the other group (Misinformed Old) received ‘new’ cues prior to 30 (out of 60 total) old-item trials. The critical comparisons were the identification speeds for items correctly cued relative to those incorrectly cued.

Thirty two UCL students (15 male) with an overall mean age of 22.9 years (*SD* = 3.6) participated in this experiment for a small payment. They were equally divided between the two groups. The stimuli those used in Experiment 3b, and in total there were 60 study trials and 120 CID trials (60 old and 60 new).

Results and Discussion

Study phase performance (categorisation accuracy) did not significantly differ between groups ($M = 95.8\%$; $SD = 3.9$, and $M = 97.9\%$; $SD = 4.7$, for the Misinformed New and Misinformed Old groups, respectively), $t(30) = 1.38$, $p = .18$.

RTs were recorded for items for which participants received correct information (correctly-cued old and correctly-cued new items) and those for which they received incorrect information (incorrectly-cued new items in the Misinformed New group and incorrectly-cued old items in the Misinformed Old group; see Figures 5A and B). One-way repeated measured ANOVAs indicated a significant main effect of item in the Misinformed New group, $F(2, 30) = 29.3$, $p < .001$, $\eta_p^2 = .66$, and the Misinformed Old group, $F(2, 30) = 5.67$, $p = .01$, $\eta_p^2 = .28$. In the Misinformed New group, RTs for incorrectly-cued new items (new items cued ‘old’) and correctly-cued new items did not significantly differ, $t(15) = 0.82$, $p = .43$, and in the Misinformed Old group, RTs for incorrectly-cued old items (old items cued ‘new’) and correctly-cued old items did not significantly differ, $t(15) = 0.28$, $p = .78$. In the Misinformed New group, correctly-cued old items were identified significantly faster than correctly-cued new items, $t(15) = 7.66$, $p < .001$, $d = 0.69$ (a priming effect), and incorrectly-cued new items, $t(15) = 4.53$, $p < .001$, $d = 0.55$, and in the Misinformed Old group, correctly-cued new items were identified significantly slower than correctly-cued old items, $t(15) = 2.47$, $p = .03$, $d = 0.34$ (a priming effect), and incorrectly-cued old items, $t(15) = 2.44$, $p = .03$, $d = 0.36$. In other words, identification speed only varied in each group as a function of the actual status of the items (old/new) and not the information provided to participants about the items. Collapsed across true item status, priming (Figure 5C) did not significantly differ between groups, $t(30) = 1.06$, $p = .30$.

In sum, providing participants with explicit information (correct or incorrect) does not affect the speed of identification of items in the CID task, nor the magnitude of priming. Even with a cue that should have created a small search set in explicit memory (Experiment 3b), no benefit to identification speed was observed.

General Discussion

We investigated whether implicit memory is preserved in normal aging despite reduced explicit memory. To overcome the problems associated with measuring these memory phenomena in separate experimental phases, each test item was appraised for a recognition and priming judgement in a single trial (Experiment 1). Recognition memory was significantly lower in older relative to young adults, and priming was numerically reduced. These observations were replicated in Experiment 2 when the priming and recognition tasks were presented separately to reduce test awareness. In both experiments the priming task was statistically less reliable than recognition, thus it was more difficult to detect effects in priming. However, the age difference in priming reached significance in a combined analysis of our data which increased statistical power. The findings present a challenge to the notion of multiple memory systems, as the age-related dissociation between recognition and priming is often cited as evidence for independent systems. Moreover, the Berry et al. single-system model provided an excellent fit to the data of Experiment 1.

Other studies have observed an age effect in priming, but this is often attributed to explicit contamination. Explicit processing during an implicit test could feasibly contribute to an age effect in priming, because young and older adults may differ in their ability to use such a strategy. Russo and Parkin (1993) found that priming differed with age on a fragmented picture completion task, but showed that the effect disappeared when explicit memory was equated between groups (young participants performed a dual study task). Geraci and Barnhardt (2010) found greater levels of test awareness and priming in young relative to older adults on word-stem completion and category production tasks, and a stronger relationship between the two, which was taken as evidence that awareness mediates age effects in priming. Furthermore, Park and Shaw (1992) demonstrated a small, non-significant age difference in priming on a word-stem completion task, but means were identical for unaware young and older participants (.08). In contrast, in Experiment 2 of the present study the numerical age difference in priming persisted when the data from aware participants were removed. Further, in Experiment 3 we showed that priming in the CID task is generally unaffected by explicit processing, thus it is unlikely that the age difference in priming was mediated by explicit contamination.

Several other studies have examined the effect of test awareness and explicit processing on priming in young individuals, but results have been mixed. Bowers and Schacter (1990) found no difference in priming between informed and uninformed participants on a word-stem completion task, and no interaction with a levels-of-processing manipulation, however priming was greater in uninformed participants who spontaneously became test aware relative to those who did not (Experiment 1). In contrast, Mace (2003) reported enhanced priming in informed relative to uninformed participants for items studied under semantic but not nonsemantic conditions. Brown, Jones, and Mitchell (1996) found no difference in priming when identification and recognition judgements were presented concurrently on every trial relative to separate experimental phases, suggesting that the presence of the recognition task alongside priming, which arguably induces greater test awareness than when the tasks are separated, does not affect priming (see also Stark and McClelland, 2000). The present Experiment 3 offers a more robust test of the claim that explicit processing affects priming, as providing participants with information about the item's status before each trial creates optimal conditions for the successful use of an explicit strategy. No difference in priming was observed under these conditions in comparison to participants who performed the CID task under implicit instructions (and did not spontaneously become test aware). A similar approach was used by Brown, Nesblett, Jones, and Mitchell (1991), who found no difference in priming on picture and word naming tasks between participants who witnessed old and new trials in separate blocks at test and were informed which block contained which type of item, versus a group who were uninformed and witnessed interspersed old and new trials within the test.

Some may argue that explicit processing may have hindered the performance of older individuals in the CID task, resulting in lower priming in this group. This is unlikely given the evidence that the CID task is unaffected by both optimal and adverse explicit processing: performance was not improved when correct explicit information was provided to support performance (Experiments 3a and 3b), and it was not worsened when explicit processing was disadvantageous because incorrect cues were provided (Experiment 3c). Thus, although some priming tasks may be more susceptible to the effects of explicit contamination, we suggest that this does not pose a concern in the CID-R task. As MacLeod (2008) pointed out, this may be because identification is usually accomplished too quickly for the engagement of intentional memory strategies.

The current study revealed an effect of varying the number of old trials in the priming task (Experiment 3a). A high proportion of old trials resulted in stronger priming relative to a low proportion. A handful of prior studies have also varied the ratio of old/new test items, but because this has typically been done to bolster an instructional manipulation (i.e., uninformed participants are exposed to fewer old trials), it is impossible to unravel the differential contributions of the factors to outcomes. Jacoby (1983) reported enhanced priming on a word naming task in informed participants who witnessed 90% old trials at test relative to uninformed participants who were exposed to 10% old trials (see also Richardson-Klavehn, Lee, Joubran, & Bjork, 1994). For reasons noted in the discussion of Experiment 3a we believe that the present effect of proportion on priming is not due to test awareness. Rather, it could be the case that a disproportionate number of old and new trials creates a motivational imbalance which affects response speeds. Identification times were faster overall in the High conditions, suggesting that participants' general ability to identify objects was boosted under such circumstances. Overall, the influence on priming of varying the ratio of old/new test trials is not well understood and merits further investigation.

The present investigation leads to the conclusion that the age-related reduction in priming is genuinely due to a decline in memory function rather than explicit contamination in the priming task. The issue of 'implicit contamination' is also interesting – that is, the extent to which fluent processing due to prior exposure to test stimuli affects performance on explicit tasks. Generally, it is believed that the contribution of fluency from priming to recognition memory in the CID-R paradigm is very minor (Conroy et al., 2005; Sheldon & Moscovitch, 2010), but future research could explore potential interactions with age effects in explicit memory.

Finally, the delay effects in Experiment 1 deserve some consideration. We observed a weak delay effect on recognition in both groups (significant only at the one-tailed level). Priming was unaffected by the delay, but the small effect size coupled with the lower reliability of the CID task renders this unsurprising. The numerical priming patterns mirrored those in recognition, meaning it would be difficult to interpret this finding as support for multiple-systems. Furthermore, the single-system model predicted a very small change in priming across the delay.

To conclude, we provide evidence that priming and recognition are reduced by normal aging on a task that is immune to explicit contamination. This suggests that normal age-related memory decline leads to the compromise of a single system which supports both explicit and implicit expressions of memory, and accordingly our formal model fits the data closely.

References

Abbenhuis, M. A., Raaijmakers, W. G. M., Raaijmakers, J. G. W., & Van Woerden, G. J. M. (1990). Episodic memory in dementia of the Alzheimer type and in normal aging: Similar impairment in automatic processing. *The Quarterly Journal of Experimental Psychology, 42*, 569–583.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Caski (Eds.), *Second international symposium on information theory*, pp. 267–281. Budapest: Academiai Kiado.

Berry, C. J., Henson, R. N. A., & Shanks, D. R. (2006). On the relationship between repetition priming and recognition memory: Insights from a computational model. *Journal of Memory and Language, 55,* 515–533.

Berry, C. J., Shanks, D. R., & Henson, R. N. A. (2008a). A single-system account of the relationship between priming, recognition, and fluency. *Journal of Experimental Psychology*: *Learning, Memory, and Cognition, 34,* 97–111.

Berry, C. J., Shanks, D. R., & Henson, R. N. A. (2008b). A unitary signal-detection model of implicit/explicit memory. *Trends in Cognitive Sciences*, *12*, 367–373.

Berry, C. J., Shanks, D. R., Li, S., Rains, L. S., & Henson, R. N. A. (2010).Can “pure” implicit memory be isolated? A test of a single-system model of recognition and repetition priming. *Canadian Journal of Experimental Psychology, 64,* 241–255.

Berry, C. J., Shanks, D. R., Speekenbrink, M., & Henson, R. N. A. (2012). Models of recognition, repetition priming, and fluency: Exploring a new framework. *Psychological Review, 119,* 40–79.

Bowers, J. S., & Schacter, D. L. (1990). Implicit memory and test awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 404–416.

Brown, A. S., Jones, T. C., & Mitchell, D. B. (1996). Single and multiple test repetition priming in implicit memory. *Memory*, *4*, 159–173.

Brown, A. S., Neblett, D. R., Jones, T. C., & Mitchell, D. B. (1991). Transfer of processing in repetition priming: Some inappropriate findings. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 514–525.

Buchner, A., & Brand, M. (2003). Further evidence for systematic reliability differences between explicit and implicit memory tests. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 56*, 193–209.

Buchner, A., & Wippich, W. (2000). On the reliability of implicit and explicit memory measures. *Cognitive Psychology, 40*, 227–259.

Butler, B. C., & Klein, R. (2009). Inattentional blindness for ignored words: Comparison of explicit and implicit tasks. *Consciousness and Cognition, 18,* 811–819.

Chiarello, C, & Hoyer, W. J. (1988). Adult age differences in implicit and explicit memory: Time course and encoding effects. *Psychology and Aging, 3,* 358–366.

Conroy, M. A., Hopkins, R. O., & Squire, L. R. (2005). On the contribution of perceptual fluency and priming to recognition memory. *Cognitive, Affective, and Behavioural Neuroscience, 5,* 14–20.

Fleischman, D. A. (2007). Repetition priming in aging and Alzheimer’s disease: An integrative review and future directions. *Cortex, 43,* 889–897.

Gabrieli, J. D. E. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology, 49,* 87–115.

Geraci, L., & Barnhardt, T. M. (2010). Aging and implicit memory: Examining the contribution of test awareness. *Consciousness and Cognition, 19,* 606–616.

Hultsch, D. F., Masson, M. E. J., & Small, B. J. (1991). Adult age differences in direct and indirect tests of memory. *Journal of Gerontology: Psychological Sciences, 46*, 22–30.

Jacoby, L. L. (1983). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory and Cognition, 9,* 21–38.

La Voie, D., & Light, L. L. (1994). Adult age differences in repetition priming: A meta-analysis. *Psychology and Aging, 9,* 539–553.

Kinder, A., & Shanks, D. R. (2001). Amnesia and the declarative/nondeclarative distinction: A recurrent network model of classification, recognition, and repetition priming. *Journal of Cognitive Neuroscience, 13,* 648–669.

Kinder, A., & Shanks, D. R. (2003). Neuropsychological dissociations between priming and recognition: A single-system connectionist account. *Psychological Review, 110,* 728–744.

Light, L. L., La Voie, D., Valencia-Laver, D., Albertson Owens, S. A., & Mead, G. (1992). Direct and indirect measures of memory for modality in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1284–1297.

Light, L. L., & Singh, A. (1987). Implicit and explicit memory in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 531–541.

Mace, J. H. (2003). Study-test awareness can enhance priming on an implicit memory task: Evidence from a word completion task. *The American Journal of Psychology, 116,* 257–279.

MacLeod, C. M. (2008). Implicit memory tests: Techniques for reducing conscious intrusion. In J. Dunlosky & R. A. Bjork (Eds.). *Handbook of memory and metamemory* (245-263). New York. Psychology Press.

Meier, B., & Perrig, W. J. (2000). Low reliability of perceptual priming: Its impact on experimental and individual difference findings. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53A,* 211–233.

Mitchell, D. B., Brown, A. S., & Murphy, D. R. (1990). Dissociations between procedural and episodic memory: effects of time and aging. *Psychology and Aging, 5,* 264–276.

Mitchell, D. B., & Bruss, P. J. (2003). Age differences in implicit memory: Conceptual, perceptual, or methodological? *Psychology and Aging, 18,* 807–822.

Ostergaard, A. L. (1998). The effects on priming of word frequency, number of repetitions, and delay depend on the magnitude of priming. *Memory and Cognition, 26,* 40–60.

Park, D. C., & Shaw, R. J. (1992). Effect of environmental support on implicit and explicit memory in younger and older adults. *Psychology and Aging, 7,* 632–642.

Richardson-Klavehn, A., Lee, M. G., Joubran, R., & Bjork, R. A. (1994). Intention and awareness in perceptual identification priming. *Memory & Cognition, 22,* 293–312.

Roediger, H. L., & McDermott, K. B. (1993). Implicit memory in normal human subjects. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 8, pp. 63–131). Amsterdam, the Netherlands: Elsevier.

Russo, R., & Parkin, A. J. (1993). Age differences in implicit memory: More apparent than real. *Memory and Cognition, 21,* 73–80.

Schacter, D. L., Cooper, L. A., & Valdiserri, M. (1992). Implicit and explicit memory for novel visual objects in older and younger adults. *Psychology and Aging, 2,* 299–308.

Schneider, J. (2002). Near vision test card. Retrieved from http://www.i-see.org/block letter eye chart.pdf.

Shanks, D. R., & Perruchet, P. (2002). Dissociation between priming and recognition in the expression of sequential knowledge. *Psychonomic Bulletin & Review, 9,* 362–367.

Shanks, D. R., Wilkinson, L., & Channon, S. (2003). Relationship between priming and recognition in deterministic and probabilistic sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 248–261.

Sheldon, S., & Moscovitch, M. (2010). Recollective performance advantages for implicit memory tasks. *Memory, 18,* 681–697.

Silver, N. C., & Dunlap, W. P. (1987). Averaging correlation coefficients: Should Fisher's z transformation be used? *Journal of Applied Psychology, 72,* 146–148.

Soldan, A., Hilton, J. H., Cooper, L. A., & Stern, Y. (2009). Priming of familiar and unfamiliar visual objects over delays in young and older Adults. *Psychology and Aging, 24*, 93–104.

Squire, L. R. (1994). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. In D. L. Schacter & E. Tulving (Eds.), *Memory systems* (pp. 203–231). Cambridge, MA: MIT Press.

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology Learning and Memory, 82*, 171–177.

Squire, L. R. (2009). Memory and brain systems: 1969–2009. *Journal of Neuroscience, 29,* 12711–12716.

Stark, C. E. L., & McClelland, J. L. (2000). Repetition priming of words, pseudowords, and nonwords. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 945–972.

Sullivan, M. P., Faust, M. E., & Balota, D. A. (1995). Identity negative priming in older adults and individuals with dementia of the Alzheimer type. *Neuropsychology; 9,* 537–555.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247,* 301–306.

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language, 50*, 289–335.

Footnotes

[1] In Experiments 1 and 2, young and older adults differed in visual acuity, vocabulary, and processing speed scores (see Tables 1 and 3). Vocabulary and processing speed scores also differed between groups in the additional experiment used in the pooled data analysis (Table 4). Scores on these tests were not significantly correlated with priming (the largest correlation was between priming for items studied 60 minutes before testing and processing speed scores for young adults in the additional experiment used in the pooled data-analysis, $r = .33$, $p = .15$), thus it is unlikely that these differences contributed towards the age-difference in priming.

[2] In the CID-R task, the recognition judgement immediately follows the identification (CID) segment of the trial. In Experiment 1 we only included in the analysis recognition judgements that were associated with valid identification trials (i.e., trials in which the item was correctly identified with a latency greater than 200 ms and less than 3 SD from the mean). This is because the single-system model which was used to model the experiment is fitted to pairs of recognition/identification data points. The procedure resulted in a loss of 368 of the total 4800 trials (7.67%) across participants.

[3] Priming = (RTnew – RTold) / RTnew. A proportional priming score was used as baseline (new item) identification times were slower in older relative to young adults in all experiments (all t's > 2, p's < .05). Furthermore, a pure difference priming score was correlated with baseline RTs in Experiment 1 (Immediate items: $r = .44$, $p = .005$; Delayed items: $r = .57$, $p < .001$), whereas the proportional score largely overcame this (Immediate items: $r = .20$, $p = .22$; Delayed items: $r = .35$, $p = .03$). For all experiments except Experiment 3a, priming was calculated using median identification RTs, but the qualitative pattern of results was identical when means were used. In Experiment 3a analyses were based on mean identification RTs due to unequal numbers of old and new items in the test phase.

[4] Post-hoc power analyses demonstrated that the power of our individual experiments to detect an age difference in priming was low (ranging from .17 to .26 based on actual η_p^2 values ranging from 0.029 to 0.045). This is also the case in many other studies (an excellent discussion is provided by LaVoie and Light, 1994).

[5] The data from Experiment 1 were most suitable for modeling because an identification RT plus recognition judgment was captured concurrently for each item at test. This is more constraining for the model relative to when identification times and recognition judgements are captured using different items and in distinct experimental phases (as in Experiment 2).

Appendix

Applying the Single-System Model to Experiment 1

A full description of the single-system model and data fitting procedures is given in Berry et al. (2012). A brief overview is given here. The model assumes that each item at test is associated with a strength-of-evidence value, f, which is a normally distributed, random variable with mean μ and standard deviation σ_f. Because old items were presented previously in the experiment, they will tend to have a greater value of f than new items. Hence, μ_{old}, the mean value of f for old items, is typically greater than that of new items (the value of which is fixed to zero). To generate a recognition response for an item, a random noise value, e_{r}, is first sampled and added to f to give J_{r}. e_{r} is specific to recognition, and is normally distributed with a mean of zero and standard deviation of σ_{r} (which is typically fixed to equal σ_f). Each item's value of J_{r} is then compared with a criterion, C. If an item's value of J_{r} exceeds C, the item will be judged old, otherwise it will be judged new (as in standard signal detection theory).

Each item's value of f is also used to calculate the item's priming measure (this is what makes the model a single-system model). Importantly, f is combined with a different, independent source of noise, and is scaled to produce an identification RT. Greater values of f are assumed to produce shorter identification RTs:

$$\text{RT} = b - sf + e_{\text{p}} \qquad \text{(A1)}$$

where b is the RT-intercept (the expected RT for new items under the model), s is a scaling parameter that determines the rate of change in RT with f, and e_{p} is a random sample of noise that is specific to the priming task, is normally distributed, has a mean of zero and standard deviation of σ_{p}. Importantly, f, e_{p}, and e_{r} are uncorrelated with one another. The noise parameters represent the influence of non-memorial factors upon performance in each task.

Such factors could include trial-to-trial variability in the placement of the decision criterion (e.g., in the case of e_r), or the amount of perceptual information available from a stimulus at test (e.g., in the case of e_p; see e.g., Ostergaard, 1998). We have previously assumed that the variance of the noise associated with the priming task is typically greater than that of the recognition task (where the same response metric is used for recognition and priming, and the variances are therefore directly comparable). This assumption is strongly supported by the greater reliability of recognition compared to priming observed in the data shown in Table 6.

A consequence of modeling recognition and priming in this manner is that changes in μ_{old} will tend to produce changes in both recognition and priming. However, because of the different ways in which recognition and priming measures are generated, and because of the greater variance of the noise typically associated with priming, the effect of changing μ_{old} will tend to be greater for recognition than priming.

The model was fit to each individual's data in Experiment 1, using the data from each trial at test. Immediate and delayed items were modeled by assuming that there were separate distributions of *f* for each type of item. There were six free parameters: These were $\mu_{Immediate}$, the mean *f* of Immediate items; μ_{Delay}, the mean *f* of delayed items; *b*, the RT intercept; *s* the rate of change of RT with *f*; σ_p, the variance of the noise associated with the priming measure, and *C*, the old/new recognition judgment criterion. Certain parameter values were fixed (as in Berry et al., 2012): $\sigma_f = \sigma_r = \sqrt{0.5}$, and $\mu_{New} = 0$. Parameters were estimated for each individual.

The likelihood for a pair of observations, *Z* and RT, where *Z* denotes the recognition judgment on a given trial and RT denotes the identification RT, is given as

$$L(Z, \mathrm{RT}|\mathrm{I}) = \left[\Phi\left(C_j \middle| \mu_{jr|\mathrm{RT,I}}, \sigma^2_{jr|\mathrm{RT}}\right) - \Phi\left(C_{j-1} \middle| \mu_{jr|\mathrm{RT,I}}, \sigma^2_{jr|\mathrm{RT}}\right)\right] \times \phi(\mathrm{RT}|\, b - s\mu_{\mathrm{I}}, \sigma^2_{\mathrm{RT}}) \quad \text{(A2)}$$

where I is the item type and I = new, immediate, delayed; Φ is the cumulative normal distribution function; ϕ is the normal density function; and $\sigma^2_{RT} = s^2\sigma^2_f + \sigma^2_p$ (from Equation A1). j = 1 when Z = "new", and j = 2 when Z = "old"; $C_0 = -\infty$, $C_1 = C$ and $C_2 = \infty$. $\mu_{Jr|RT,I}$ and $\sigma^2_{Jr|RT}$ are the mean and variance of the conditional distribution of J_r given RT, and are calculated as follows:

$$\mu_{Jr|RT,I} = \mu_I - \frac{s\sigma^2_f(RT - b + s\mu_I)}{s^2\sigma^2_f + \sigma^2_p}$$

and

$$\sigma^2_{Jr|RT} = \sigma^2_f + \sigma^2_r - \frac{s^2\sigma^4_f}{s^2\sigma^2_f + \sigma^2_p},$$

where $\mu_{new} = 0$ when I = new; $\mu_{immediate} \geq 0$ when I = immediate; and $\mu_{delayed} \geq 0$ when I = delayed.

An optimisation routine was used to find the parameter values that maximised the log likelihood of the model given the data. The summed log likelihoods across trials and participants are given in Table A1 for each group, and the maximum likelihood parameter estimates are shown in Table A2. It is evident that the estimates of both $\mu_{Immediate}$ and μ_{Delay} are lower in the older group than the younger group, and in both groups, $\mu_{Immediate}$ is greater than μ_{Delay}. This suggests that the memory signal driving recognition and priming was weaker in the older group than the young group, and that the memory signal was weaker for delayed items than immediate items. The criterion C is approximately equal in each group. The estimates of b, s, and σ_p were all greater for the older group than the younger group, which may reflect the generally longer identification RTs in the older group.

The expected values for recognition and priming for each participant were calculated using the formulae in Table A3.

Table A1

Goodness of fit of the SS model.

Group	*p*	ln(L)	AIC
Older	6	-18612.85	37465.71
Young	6	-19332.11	38904.22

Note. AIC = Akaike (1973) criterion, calculated as AIC = −2ln(L) + 2*P*, where $P = p \times n$ is the total number of free parameters for each fit, where *p* is the number of free parameters for each model, and *n* is the number of participants modeled in each group ($n = 20$). The total number of data points fit was as follows: Older = 2117, Young= 2315.

Table A2

Mean and Standard Deviation (in Parenthesis) of the Parameter Estimates of the SS model.

Parameter	Older	Young
$\mu_{\text{Immediate}}$	1.10 (0.45)	1.50 (0.76)
μ_{Delayed}	0.95 (0.50)	1.31 (0.73)
C	0.77 (0.43)	0.79 (0.64)
b	2189 (635)	1470 (442)
s	199 (166)	140 (100)
σ_p	936 (315)	630 (197)

Table A3

Expected values for measures of recognition and priming in the single-system model.

Measure	Model expected value
P(Hit| Immediate)	$1 - \Phi(C - \mu_{\text{Immediate}})$
P(Hit| Delay)	$1 - \Phi(C - \mu_{\text{Delay}})$
P(False Alarm)	$1 - \Phi(C)$
$d'_{\text{Immediate}}$	$\mu_{\text{Immediate}}$
d'_{Delayed}	μ_{Delayed}
E[RT| new]	b
E[RT| Immediate]	$b - s\mu_{\text{Immediate}}$
E[RT| Delayed]	$b - s\mu_{\text{Delayed}}$
Proportional priming measure (Immediate items)	$s\mu_{\text{Immediate}} / b$
Proportional priming measure (Delayed items)	$s\mu_{\text{Delayed}} / b$

Table 1

Participant characteristics for Experiment 1

	Young (n = 20)	Older (n = 20)
	M (SD)	*M (SD)*
Age (years)	24.3 (3.8)	69.1 (5.5)
Education (years)	16.5 (1.2)	15.9 (2.5)
Visual acuity *	27.0 (5.5)	32.6 (6.1)
WAIS-III Vocabulary *	54.9 (10.2)	64.7 (2.2)
WAIS-III Digit Symbol (processing speed) *	89.9 (14.0)	66.4 (15.8)
Wechsler Test of Adult Reading (WTAR)	46.9 (3.8)	48.8 (2.5)
Mini Mental State Exam (MMSE)	-	29.8 (0.6)

Note. Visual acuity measured using the Near Vision Test Card (Schneider, 2002), viewed at a distance of 16 inches whilst wearing corrective glasses. Participants indicated the smallest set of letters that they could comfortably read, and scores on this test can range from 16 (highest acuity) to 160 (lowest acuity). The WAIS-III (Wechsler Adult Intelligence Scale III) subtests Vocabulary and Digit Symbol Substitution have maximum scores of 66 and 133, respectively, and the maximum score on the WTAR is 50. The maximum score on the MMSE is 30. A score of 23 or lower indicates probable cognitive impairment, however no participants in the experiments reported here scored below 27.

* Significant difference between groups, $p < .05$

Table 2

Mean of median identification RTs for old and new items in the Young and Older adult groups in Experiments 1 and 2

	Young			Older		
	Old (immediate)	Old (delayed)	New	Old (immediate)	Old (delayed)	New
Experiment 1	1116 (310)	1110 (262)	1305 (471)	1807 (598)	1784 (541)	1962 (650)
Experiment 2	1174 (349)	1221 (366)	1405 (462)	2278 (857)	2322 (372)	2471 (785)

Note. Standard deviations in parenthesis. Collapsed across young and older individuals, baseline (new item) RTs did not significantly differ between experiments, $t(74) = 1.78$, $p = .08$.

Table 3

Participant characteristics for Experiment 2

	Young (n = 18)	Older (n = 18)
	M (SD)	*M (SD)*
Age (years)	23.9 (3.5)	73.3 (6.7)
Education (years)	16.0 (1.6)	15.8 (1.4)
Visual acuity *	29.4 (6.9)	37.8 (8.1)
WAIS-III Vocabulary *	56.6 (8.7)	65.4 (1.1)
WAIS-III Digit Symbol *	80.5 (17.9)	64.7 (18.0)
WTAR	42.7 (5.1)	49.3 (0.7)
MMSE	-	29.7 (0.5)

* Significant difference between groups, $p < .05$

Table 4

Participant characteristics for the additional experiment included in the pooled data-analysis

	Young (n = 20)	Older (n = 20)
	M (SD)	*M (SD)*
Age (years)	23.4 (3.2)	65.8 (5.5)
Education (years)	15.3 (2.6)	17.4 (4.5)
Visual acuity	31.8 (8.1)	34.4 (7.4)
WAIS-III Vocabulary *	48.0 (3.4)	61.7 (8.7)
WAIS-III Digit Symbol (processing speed) *	74.3 (14.4)	55.0 (11.6)
Wechsler Test of Adult Reading (WTAR)	44.7 (3.2)	46.8 (1.7)
Mini Mental State Exam (MMSE)	-	28.4 (1.2)

* Significant difference between groups, $p < .05$

Table 5

Recognition and priming scores for Young and Older adults in the additional experiment included in the pooled data-analysis

	Young (n = 20)		Older (n = 20)	
	Immediate Test	Delayed Test	Immediate Test	Delayed Test
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
d′	1.69 (.89)	1.38 (.92)	1.17 (.71)	.87 (.59)
Proportion priming	.24 (.13)	.18 (.15)	.14 (.14)	.19 (.11)

Note. There were significant main effects of age group, $F(1, 38) = 5.61$, $p = .02$, $\eta_p^2 = .13$, and delay, $F(1, 38) = 6.81$, $p = .01$, $\eta_p^2 = .15$, on recognition, and no interaction, $F(1, 38) = 0.002$, $p = .96$. There was no main effect of age group, $F(1, 38) = 1.35$, $p = .25$, or delay, $F(1, 38) = 0.06$, $p = .81$, on priming, but the interaction approached significance, $F(1, 38) = 4.00$, $p = .053$.

Table 6

Split-half correlations for the recognition and priming scores

Experiment 1					
	Young		Older		
	Immediate	Delayed	Immediate	Delayed	Aggregate
Recognition					
r	.69	.32	.69	.30	.49
p	< .001	.17	< .001	.20	< .001
Priming					
r	-.06	-.03	-.31	-.04	-.12
p	.81	.89	.19	.84	.35
Experiment 2					
	Young		Older		Aggregate
Recognition					
r	.54		.46		.59
p	.02		.05		< .001
	Immediate	Delayed	Immediate	Delayed	
Priming					
r	.22	.51	.09	-.23	.14
p	.38	.03	.70	.35	.25
Additional experiment (pooled data-analysis)					
	Young		Older		
	Immediate	Delayed	Immediate	Delayed	Aggregate

Recognition					
r	.63	.67	.51	.54	.62
p	.003	.001	.02	.01	< .001
Priming					
r	.21	-.06	-.30	.20	-.01
p	.39	.79	.21	.40	.90

Note. The individual data points (e.g., for young and older adults, and immediate and delayed items) were pooled in order to generate aggregate correlation coefficients (see Silver & Dunlap, 1987, for a discussion of why it may be problematic to average correlation coefficients). The aggregate *r* values were *z*-transformed to test the null hypothesis that *r*(recognition) – *r*(priming) = 0.

Table 7

Identification RTs for old and new items in Experiments 3a and 3b

	Old	New
	M (SD)	*M (SD)*
Experiment 3a		
Informed High	1479 (375)	1665 (447)
Informed Low	1652 (303)	1757 (236)
Uninformed High	1512 (292)	1677 (316)
Uninformed Low	1667 (348)	1767 (292)
Experiment 3b		
Informed	1717 (470)	1897 (469)
Uninformed	1735 (331)	1937 (346)

Note. Mean of median identification RTs are listed for Experiments 3b

Figure 1. A: Priming mask used in the identification task and examples of picture stimuli. B: Example of the CID portion of the CID-R trial. Pictures gradually clarify from a background mask and participants must identify the objects as quickly as possible. RTs captured upon an ‘Enter’ keypress.

A

B

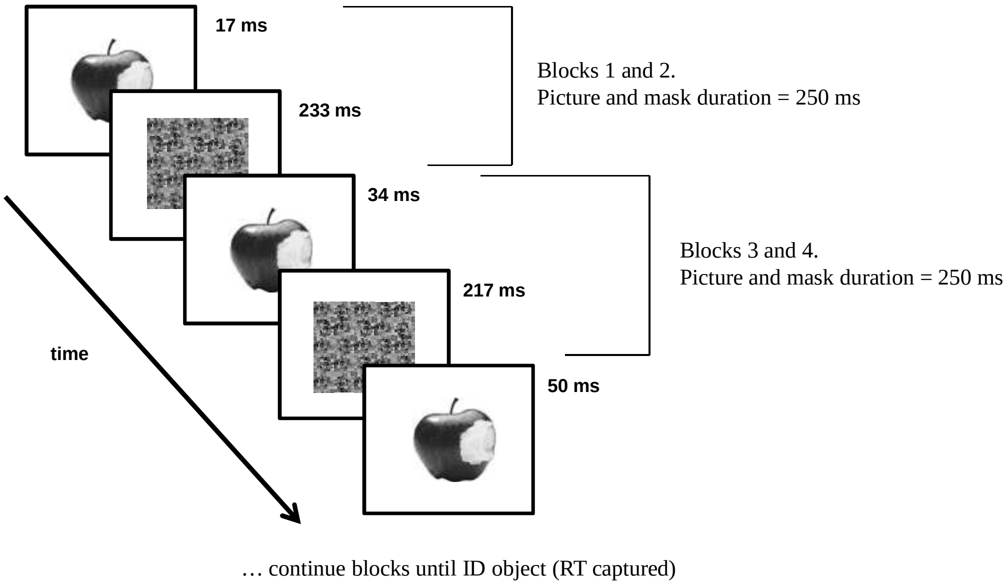


Figure 2. A: Mean *d'* scores for Young and Older adults in Experiment 1 as a function of delay. Black circles indicate single-system model fits (see appendix). B: Mean *d'* scores for Young and Older adults in Experiment 2. Error bars in indicate standard error of the mean (SEM).

A B

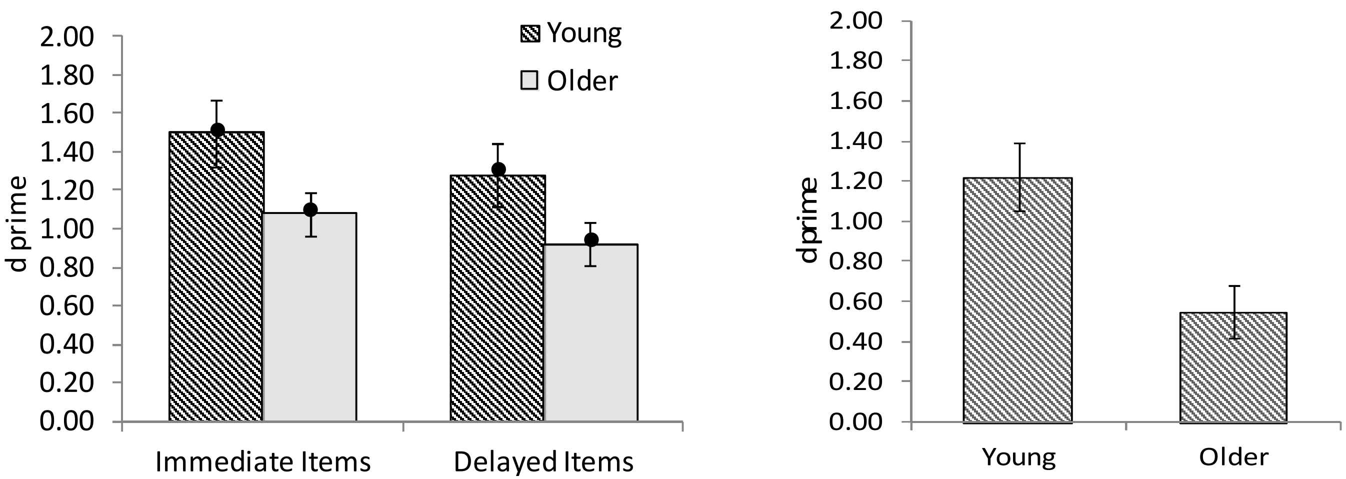
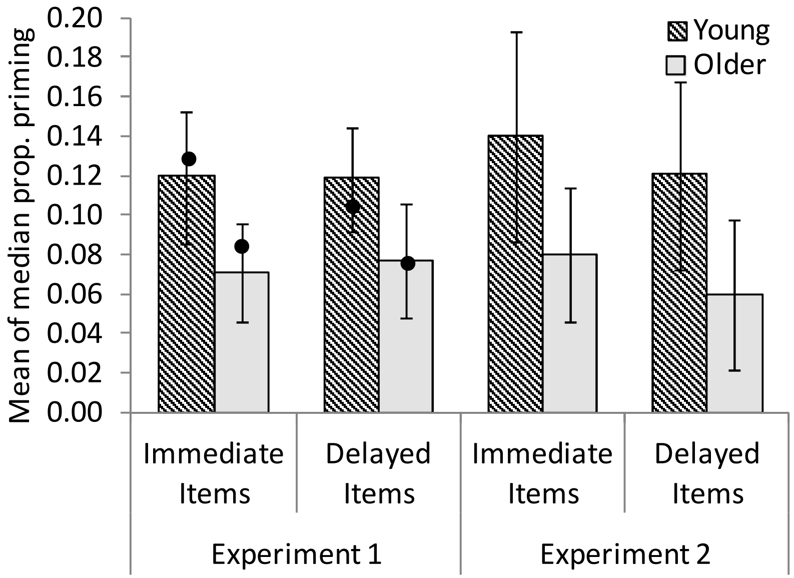


Figure 3. A: Priming in Young and Older adults as a function of delay in Experiments 1 and 2. Black circles indicate single-system model fits (see appendix). B: Pooled priming data for Young and Older adults (immediate items; *n* = 58 per group). Error bars in indicate SEM.

A



B

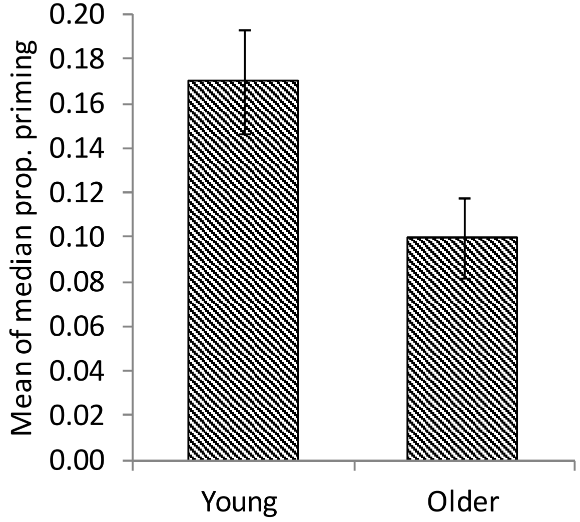


Figure 4. A: Priming in the Informed High, Informed Low, Uninformed High and Uninformed Low groups in Experiment 3a. B: Priming in the Informed and Uninformed groups in Experiment 3b. Error bars indicate SEM.

A B

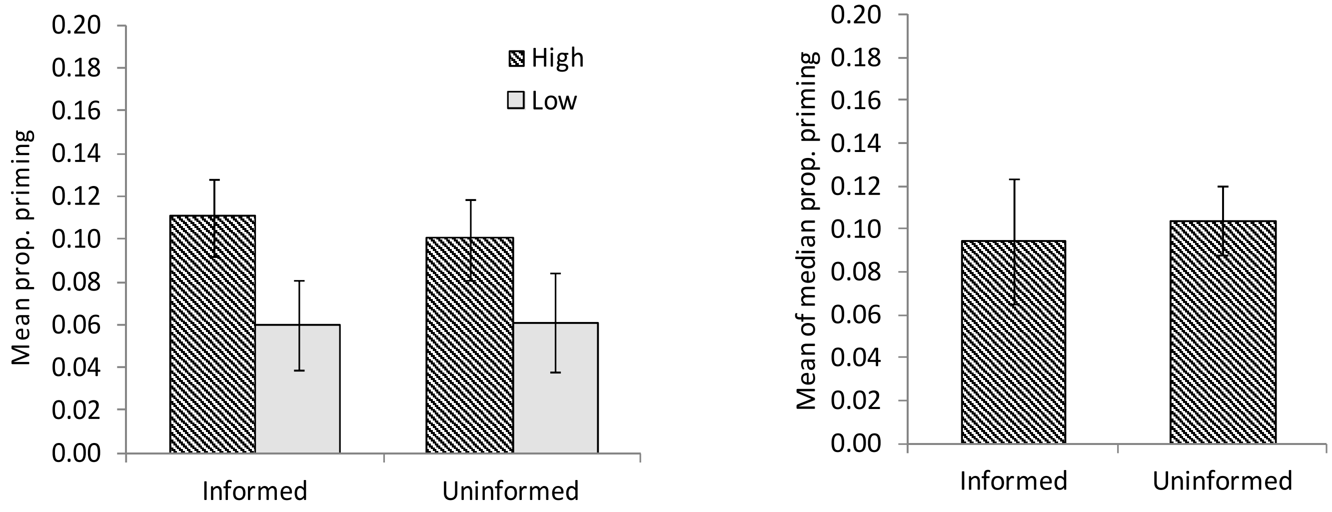


Figure 5. A: Identification times for correctly-cued old, correctly-cued new and incorrectly-cued new items in the Misinformed New group in Experiment 3c. B: Identification times for correctly-cued old, correctly-cued new and incorrectly-cued old items in the Misinformed Old group in Experiment 3c. C: Priming (calculated based on the true item status) in the Misinformed New and Misinformed Old groups in Experiment 3c. Error bars indicate SEM.

A B C